UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/971,777	10/05/2001	Stephen E. Silverman	76141.00101	2498

34661 7590 08/06/2004

CHARLES N. QUINN
FOX, ROTHCHILD, O'BRIEN & FRANKEL, LLP
2000 MARKET STREET, 10TH FLOOR
PHILADELPHIA, PA 19103

EXAMINER	
BRANT, DMITRY	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 08/06/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/971,777	SILVERMAN ET AL.	
	Examiner	Art Unit	
	Dmitry Brant	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *05 October 2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-8* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-8* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *10/5/2001*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 112

1. Claims 2,3, 6,8 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

Specifically, step F in claim 2 and step O in claim 3 recite "determining period-to-period fluctuation of fundamental frequency as the inverse of said glottal cycle for said two consecutive prominent pulses." However, specification does not support this claim. The inverse of said glottal cycle is the actual fundamental frequency. However, the fluctuation of the fundamental frequency (jitter) is not simply determined by the inverse of the glottal cycle (See Specification, "Jitter Analysis, page 17), but at least by measuring the difference between fundamental frequencies of two consecutive segments. Same analysis applies to apparatus claims 6 and 8.

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1 and 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Shubna et al.

The recitation of "method for categorizing voice samples of a person being tested for near term suicidal risk as a prelude to such testing" has not been given patentable weight because the recitation occurs in the preamble. A preamble is generally not accorded any patentable weight where it merely recites the purpose of a process or the intended use of a structure, and where the body of the claim does not depend on the preamble for completeness but, instead, the process steps or structural limitations are able to stand alone. See *In re Hirao*, 535 F.2d 67, 190 USPQ 15 (CCPA 1976) and *Kropa v. Robie*, 187 F.2d 150, 152, 88 USPQ 478, 481 (CCPA 1951).

Shubha et al. disclose:

A. setting an analysis window to a selected sample set length of 512, where the particular sample is identified as the Kth sample	Setting window length to L ms. (page 918, 2nd column, last paragraph and FIG. 1)
B. reading the Kth sample	(2nd step, FIG. 1)
C. computing wavelet transforms of such Kth sample for scales in powers of 2 running from the 1st power to the 5th	computing DyWt for scales in powers from 3rd to 5th (page 919, 1st column, 3rd paragraph)
D. storing the signal energy value as computed for each scale	inherently part of the algorithm disclosed in Fig. 1, since these values must be stored in computer memory for further comparisons.
E. checking to determine whether the Kth sample is the last of the sample set and if additional samples remain, repeating steps "b" through "d"	See Fig. 1, algorithm iterates through all segments.
F. setting the median energy distribution at the scale for 2 to the 4th power as a threshold	Threshold is set as 2 to 4th (page 919, second column, 2nd paragraph)

G. successively for each sample comparing the energy across the scales	(See FIG .1 and page 919, 1st column, last paragraph) - "... in addition to checking whether the local maxima in DyWT correlates across two scales."
J. if the segment energy at the 2 to the 4th power scale exceeds the threshold, classifying the segment as voiced otherwise classifying it as silence.	Page 919, 2nd column, 2nd paragraph.

Shubha et al. do not disclose:

A. setting an analysis window to a selected sample set length of **512**
C. computing wavelet transforms of such Kth sample for scales in powers of 2 running from **the 1st power to the 5th**
H. if the maximum energy is at the scale for 2 to the 1st power, identifying the segment as unvoiced and proceeding to the next succeeding sample
I. if the segment maximum energy is at one of the scales of 2 to the 2nd power through 2 to the 5th power, identifying the segment as being either voiced or silence

Regarding step A, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Shubha et al. to use sample set length of 512. Applicant has not disclosed whether any specific set length provides an advantage, is used for a particular purpose or solves a stated problem. One of ordinary skill in the art, furthermore, would have expected Shubha et al. to perform equally well with other sample set lengths because varying sample set length in Shubna et al. would only change processing requirements of the system, but would not affect the essence of Shubna et al's invention.

Regarding steps C, H and I, Shubna et al. do emphasize that voiced speech is usually present at 3rd to 5th scales, while unvoiced speech is only present at lower scales (p. 919, Col. 1, last paragraph). Since Shubna et al. do not try to accomplish full

voiced/unvoiced/silence determination, they only attempt to find voiced parameters at the higher scales (3^{rd} -5^{th}). However, the above disclosure is sufficient to deduce that unvoiced frequencies exist at scales in powers of 2 running from 1^{st} power to 2^{nd} power. In addition, one of ordinary skill in the art could deduce that since the voiced segments exist at or above 2 to the 4^{th} power (voice) and unvoiced segments exist at or below 2 to the 2^{nd} power, then by process of elimination, signals falling in between these thresholds must represent silent periods. Finally, Applicant's specification and other claims indicate that the disclosed invention uses only voiced segments and discards the classified silence/unvoiced segments, so there does not appear real utility from further distinguishing within the unvoiced/silence category.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Shubna et al. to perform voiced/unvoiced/silence computation based on the principles disclosed by Shubna et al. This would allow the system to distinguish between silence, voiced and unvoiced segments in order to remove silence/unvoiced segments which are not useful for fundamental frequency (pitch) calculations (p. 919, Col. 1, last paragraph) (as it is well-known in the art, unvoiced speech is not periodic and thus is not useful for pitch estimation).

3. Claims 2, 3, 6, 8 are rejected under 35 U.S.C. 103(a) as being unpatentable over Shubna et al. in view of Parson ("Voice and Speech Processing").

As per claims 2 and 6, The recitation of "method for determining jitter variations in fundamental frequency of the voice of a person being evaluated for near-term suicidal risk" has not been given patentable weight because the recitation occurs in the preamble. A preamble is generally not accorded any patentable weight where it merely recites the purpose of a process or the intended use of a structure, and where the body of the claim does not depend on the preamble for completeness but, instead, the process steps or structural limitations are able to stand alone. See *In re Hirao*, 535 F.2d 67, 190 USPQ 15 (CCPA 1976) and *Kropa v. Robie*, 187 F.2d 150, 152, 88 USPQ 478, 481 (CCPA 1951).

Shubna et al. disclose:

A. setting an analysis window to a selected sample set length of 512 where the particular sample is identified as the Kth sample	Setting window length to L ms. (page 918, 2nd column, last paragraph and FIG. 1)
B. computing the wavelet transform for the sample set at scale 2 to the 4th power, with a scale factor defined by the quotient of the wavelet center frequency at level 0 and the desired center frequency	Threshold is set as 2 to 4th (page 919, second column, 2nd paragraph)
C. selecting two consecutive segments of the vocal signal of such person which are voiced segments and generating separate pulse trains in which the heights of the pulses correspond to amplitude of positive and negative peaks of the wavelet transformed speech signal	(See FIG .1 and page 919, 1st column, last paragraph) - "... in addition to checking whether the local maxima in DyWT correlates across two scales."
D. thresholding the segments of the vocal signal to discard peaks corresponding to possible unvoiced samples	Only voiced samples are used for pitch estimation (see FIG. 1, "segment is unvoiced, set pitch period to 0" block)

E. v. taking the difference between two consecutive prominent pulses as the duration for the glottal cycle	(page 919, Col. 1, 1st paragraph)

Shubna et al. do not disclose:

A. setting an analysis window to a selected sample set length of **512**

E. computing a fundamental period over the entirety of each of the two segments by:
i. finding the location of the first peak of the autocorrelation of the smoothed spectrum to the right of the zero lag component
ii. detecting a starting pulse exhibiting the property of being larger than both the pulse immediately preceding and immediately following such pulse and being greater than 50% of the global maximum of the pulse sequence
iii. locating following prominent pulses as detected in the neighborhood of expected locations determined by the peak of the autocorrelation sequence
iv. selecting, between two sequences of positive and negative peaks, the peak having the largest magnitude
and v. taking the difference between two consecutive prominent pulses as the duration for the glottal cycle

F. determining period-to-period fluctuation of fundamental frequency as the inverse of said glottal cycle

for said two consecutive prominent pulses.

Regarding step A, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Shubha et al. to use sample set length of 512. Applicant has not disclosed whether any specific set length provides an advantage, is used for a particular purpose or solves a stated problem. One of ordinary skill in the art, furthermore, would have expected Shubha et al. to perform equally well with other sample set lengths because varying sample set length in Shubna et al. would only change processing requirements of the system, but would not affect the essence of Shubna et al's invention.

Regarding steps E and F, Shubna et al. disclose computing the fundamental frequency as the inverse of time interval between local peaks (Shubna et al., page 919, Col. 1, 1st paragraph.) The method of finding the peaks using harmonic-peak-based detection is well-known in the art. For example, "Voice and Speech Processing" by Thomas Parson describes this method on pages 205-206. Firstly, Parson teaches that filtered (smoothed) autocorrelation of signal spectrum points to the approximate location of Fo (pages 198-199). Parson also teaches that "harmonic peaks occur at integer multiples of the pitch frequency (Fo)" and that "the differences in peak frequencies are integer multiples of the pitch frequency." As a result, harmonic-peak-based detection proceeds by searching for peaks in the range of the estimated peach (from autocorrelation) and fine-tuning the results by finding the maximum peaks within these regions (see the discussion of several different methods of pages 205-206, such as Snow and Hughes).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Shubna et al. to compute the fundamental frequency using a well-known method of harmonic-peak-detection (see Parson's book) in order to fine-tune the selection of peaks and thus establish the best fundamental frequency estimate.

As per claims 3 and 8, the preamble has not been given patentable weight. Claims 3 and 8 combine elements of claim 1(7) and claim 2 (6), without adding new

limitations. Steps A-J represent the determination of the voiced segments (see claims 1, 7). Steps K-O represent pitch estimation which requires voiced samples (as required by step L). Therefore, steps A-J are rejected for the same reasons as steps A-J in claims 1 (7). Steps K-O are rejected for the same reasons as steps B-F in claim 2 (6).

4. Claims 4 and 5 are rejected under 35 U.S.C. 103(a) as being unpatentable over France et al. ("Acoustical Properties of Speech as Indicators of Depression and Suicidal Risk", published July 2000).

France et al. disclose:

A method for assessing near-term suicidal risk through voice analysis independently of verbal content of the voice, comprising:	See title of the article
eliciting a voice sample from a person to be evaluated for near- term suicidal risk and converting said sample into electronically processable signal form	(page 832, part B., second paragraph from the buttom: "Approximaately 2 min and 30 s of unedited speech ...)
time-wise dividing said signal into segments according to whether the person was silent, speaking voiced words or making unintelligible unvoiced sounds	dividing signal into segments (same reference as above)
if there are two consecutive voiced segments, measuring fundamental frequency for each of said two segments	measuring Fo for each segment (page 833, first and second paragraphs)
(**partially**) comparing the *difference in measured fundamental frequency to fundamental frequency difference data* (not disclosed) for known near-term suicidal risk persons, known depressed persons not at near-term suicidal risk and non-depressed persons from a control group, to determine whether the person is at near-term suicidal risk or is merely depressed.	Differences in Fo statistics were examined for control, major-depressed and high-risk (suicidal) groups (Table 8)

France et al. do not explicitly disclose using "difference in measured fundamental frequency for said two segments" as an indication of suicidal tendencies (claim steps D and E).

However, France et al. suggest measuring Fo range for each 20 second segment (2nd paragraph, page 833). The **range** of a set of data is the difference between the highest and lowest values in the set. Here, the 20 second segment would undoubtedly contain several (2 or 3) voiced segments, since Fo can only be measured for voiced segments and 20 seconds is too long of a duration for a single voiced segment (short of a long scream, but certainly not common in speech recorded during a therapy session). As a result, the Fo range over the 20 second window would represent the Fo difference of two segments in case where there are only two voiced segments in the said 20 second window. In addition, France et al. do discuss measuring jitter (fundament frequency difference) for Fo in their article ("II. Previous Work", page 830)

As a result, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify France et al. to use "difference in measured fundamental frequency for said two segments" (jitter) as another statistic for attempting to determine whether the person is predisposed to suicide, as the article of France et al. indirectly suggests the use of this statistic for one of ordinary skill in the art.

Conclusion

3. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Petrushin (6,353,810) teaches using voice pitch (Fo) for determination of person's emotions.

Janer et al. ("Pitch Detection and Voiced/Unvoiced Decision Algorithm based on Wavelet tranforms) teach voice/unvoiced determination using wavelets.

Fukuda et al. ("Extracting Emotion from Voice") teach using pitch for emotion detection.

Zhou et al. ("Nonlinear Feature Based Classification of Speech Under Stress") teach using pitch for stress detection.

Markel ("SIFT Algorithm for Fundamental Frequency Estimation") teaches using smoothed auto correlated spectrum for pitch estimation.

4. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dmitry Brant whose telephone number is (703) 305-8954. The examiner can normally be reached on Mon. - Fri. (8:30am - 5pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Ivars Smits can be reached on (703) 306-3011. The fax phone

number for the organization where this application or proceeding is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to Tech Center 2600 receptionist whose telephone number is (703) 305- 4700.

DB

8/3/04

Nguyen Vo
8-5-04

NGUYEN T. VO
PRIMARY EXAMINER